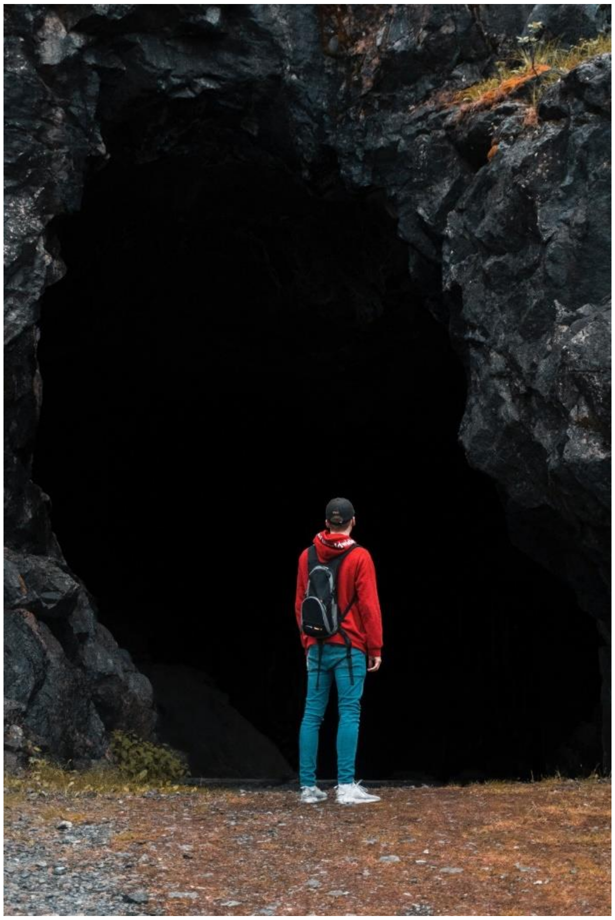
Confronting Deep Learning Systems: How Much Things Have Changed and How Much We Do Not Know

Richard Hackathorn, Bolder Technology
March 28, 2019 – Towards Data Science [link]



As companies embrace AI in their business, they are confronting the technologies of Deep Learning systems at scale. That is like entering a dark massive cave — wonderment and excitement, along with the fear that you have no clue what you are getting into!

Mostly Clueless in IT

To a colleague, I recommended the blog by Paco Nathan [1] that summarizes recent Strata surveys. The colleague is a respected expert in Business Intelligence and Data Warehousing.

His reaction was, "***How much things have changed and how much we do not know!***"
…referring to IT professionals confronting the emerging Deep Learning systems.

In my words, IT professionals are **mostly clueless** about the implications concerning AI systems being built upon existing infrastructures. Further, there are critical cultural disconnects emerging among corporate groups that must collaborate for the organization and its stakeholders to realize value from AI systems.

Although this is a disturbing commentary, I was relieved to hear this reaction. I have spent the last year trying to convince other BI/DW colleagues of the pending troubles awaiting enterprise IT. The usual reaction has been a blank stare …like they doubted my sanity!

> *TERMINOLOGY: Deep learning (DL) using artificial neural networks (ANN) is a subset of machine learning (ML). This article highlights the differences between* ***conventional machine learning*** *(without ANN) with* ***deep learning*** *(using ANN) within enterprise systems.*

So, what is the fuss about?

TL;DR — Do not enter the cave (of AI systems) until you have a map!

Pacoid Strikes Again!

Paco Nathan (or Pacoid on twitter) is self-described as "evil mad scientist" and as "player/coach". My experience leans heavily to the latter. He is a remarkable personality within the AI/ML/DS community, who is totally connected, brilliant, articulate, and humane.

Nathan collaborates with Ben Lorica, who is Chief Data Scientist at O'Reilly Media, Inc. and Program Director of Strata Data Conference and Artificial Intelligence Conference. They have surveyed the attitudes and actions of these communities, who are the early adopters — crossing (or struggling to cross) the chasm to embrace AI systems within their organizations.

So, the fuss is about paying attention to what these early adopters are telling us about current AI systems, whether they are successful, fizzle, disaster, ethically hurtful, politically naive, and everything in between.

The survey is based upon 1300 responses in mid-November 2018 with good representation by global enterprises. Hence, this information should be taken seriously. And, here is how to do so…

First, I recommend reading Pacoid's blog. [1] Second, scan the survey details in these three reports, coauthored with Lorica:

- *State of Machine Learning* (August 2018) [2]
- *Evolving Data Infrastructure* (January 2019) [3]
- *AI Adoption in the Enterprise* (February 2019) [4] — especially this one
- Also, earlier reports as cited in [5] and [6]

If you are embracing (or plan to embrace) AI systems within your organization, you will find several sobering observations. …pieces of the cave map.

Below are a few of my take-away points. **Please share your observations, in comments below.**

Take-Away: Deep Learning Has Traction

Many IT professionals to which I have spoken have the opinion that neural networks are exciting but that conventional machine learning is more than sufficient for business applications for many years to come.

However, these surveys indicate that deep learning technology has considerable traction within enterprise systems. It is no longer a darling only of researchers; this technology is being put to work!

The survey data shows that more than half are using deep learning (55%), along with reinforcement learning (22%) and transfer learning (16%), as shown in this figure.

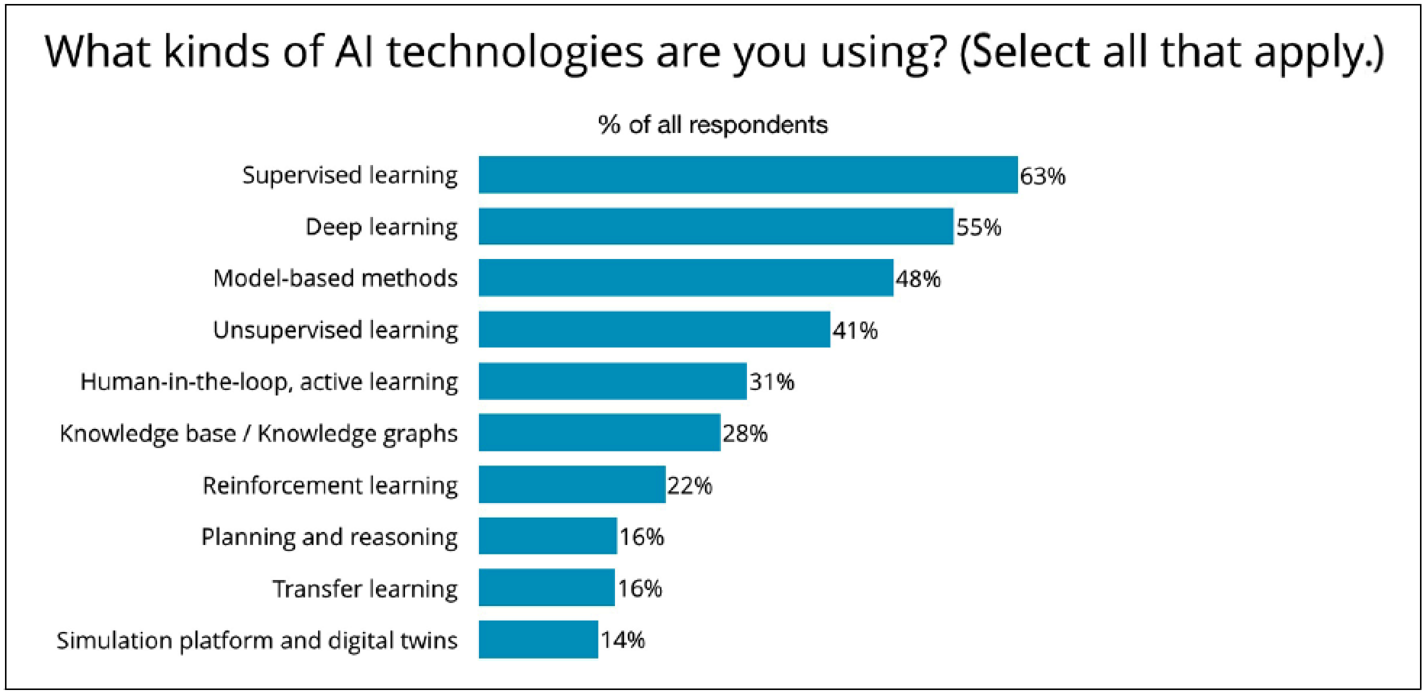


Figure 1–15. AI technologies used, from AI Adoption survey [4]

Side… One also wonders what respondents were imagining when they checked: active learning, knowledge graphs, and planning & reasoning. Need eye-to-eye interviews here!

As expected, more than half (53%) are using neural networks for image processing. However, the surprise was that significantly more are using neural networks for structured data (86%) and text (69%), as shown below.

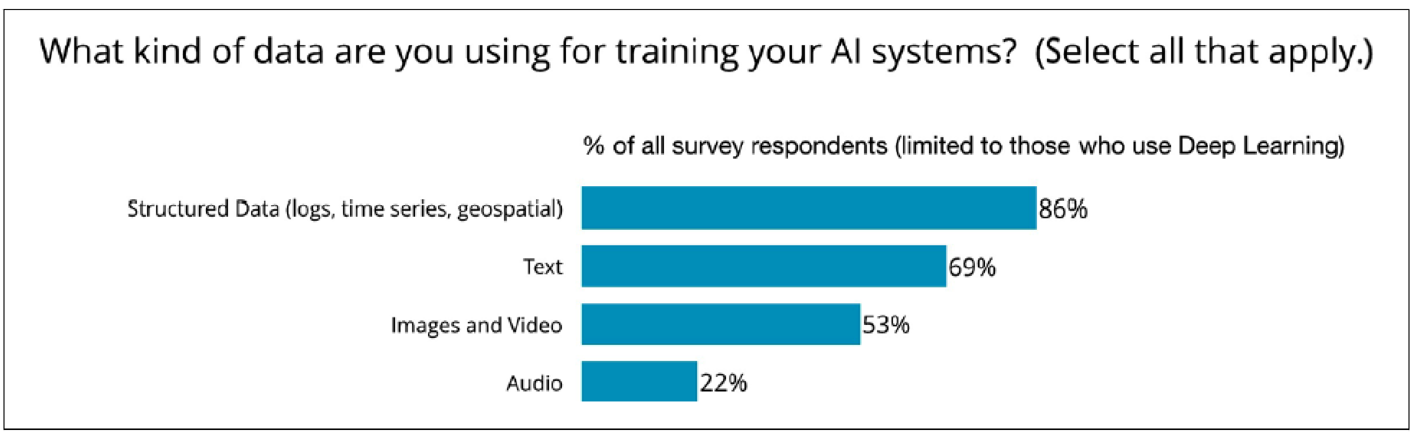


Figure 1–21. Data types, limited to deep learning respondents, from AI Adoption survey [4]

When asked about the AI tools used, a surprising number dealt primarily with neural networks, such as TensorFlow (55%), Keras (34%) and PyTorch (29%). This is significantly more than tools for conventional machine learning, such as scikit-Learn.

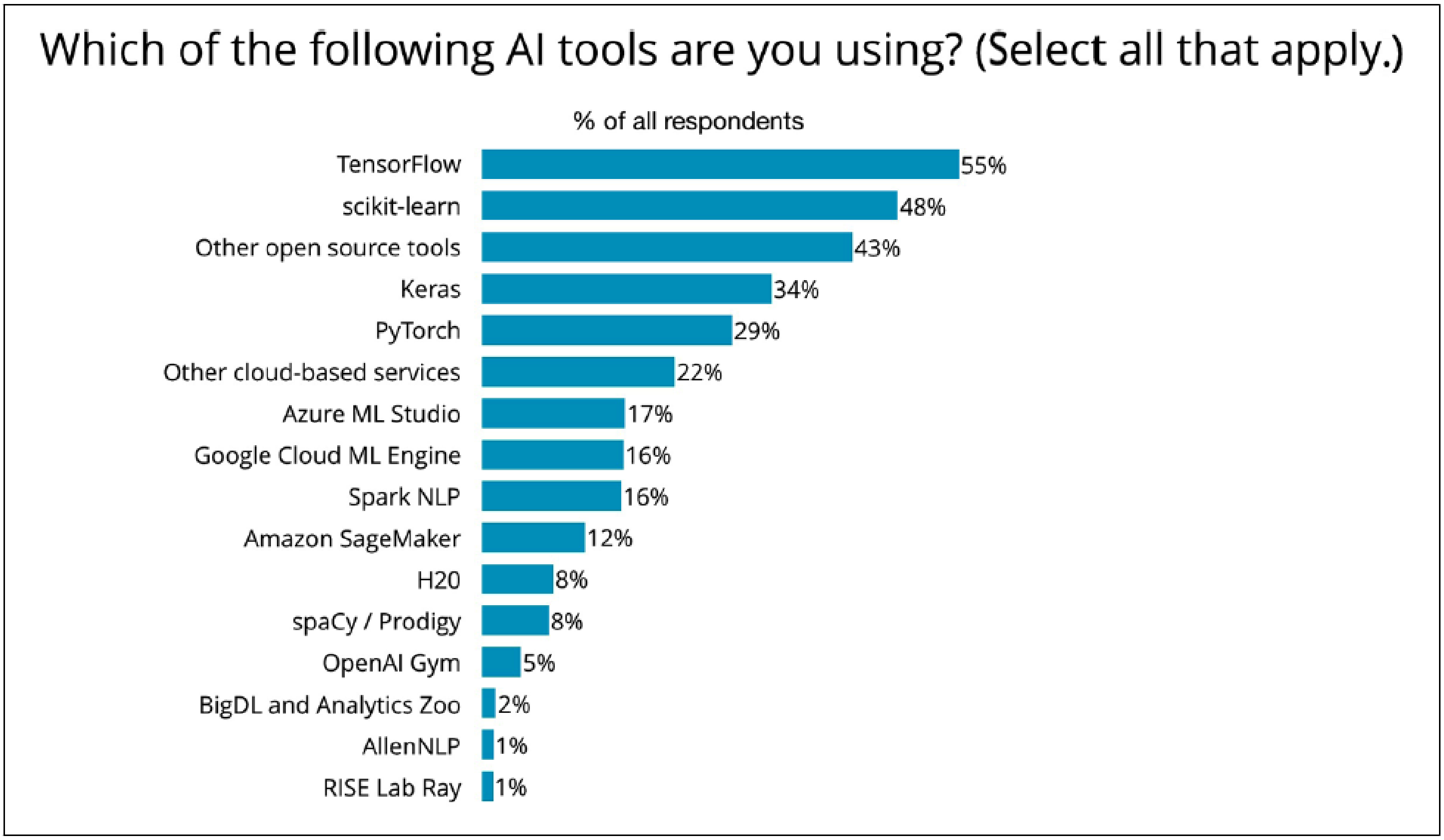


Figure 1–29. AI tools used, in [4]

Finally, subsequent figures in the AI Adoption report [4] indicate that adoption of deep learning is prevalent across industries and accelerates within companies with AI-Mature practices.

Take-Away: Keeping Up with AI Mature Folks — NOT

The classic "wait until others figure it out" strategy of IT professionals may be a disaster for some companies. **This strategy misses the rapid shift in basic paradigms that emerging AI systems are driving. Such shifts require slow cultural changes to allow the technology changes to be implemented successfully.** Let's examine this issue in detail.

Lorica and Nathan often draws distinctions among three maturity stages of AI adoption: *Not-Yet* (19%), *Evaluating* (54%), and *In-Production* or AI-Mature (27%). [4] This distinction is used to understand several factors throughout the reports.

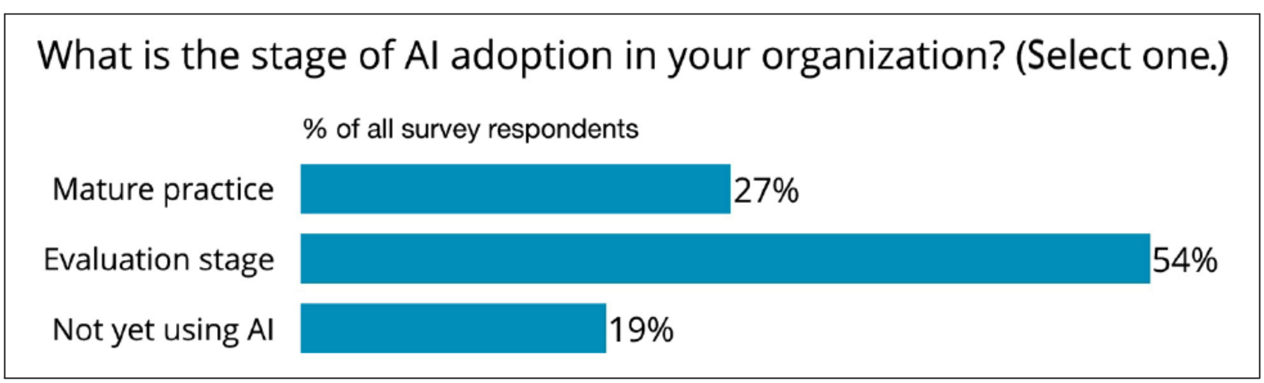


Figure 1–2. Stage of maturity for AI adoption, in [4]

Although these categories are fuzzy, observe these critical aspects. AI-Mature (In-Production) companies are constantly doing ALL three of these categories. AI-mature companies may be *Not-Yet* with a specific AI method (like reinforcement learning), but they have assigned some expert to monitor **weekly** the emerging research on that method. AI-Mature companies many be *Evaluating* many of prototypes, but they are constantly using them in challenger-champion contests with their production models. It's constant innovation at all levels, thus…

A key insight is that AI-Mature organizations are rapidly increasing the gap with the Not-Yet and Evaluating organizations, making it difficult to catch up, as shown below. Note that the bottlenecks (highlights in yellow) — culture not recognizing the need and use cases not identified — become less significant for AI-Mature companies. In contrast, lack of quality data and lack of skilled people are more significant for AI-Mature companies.

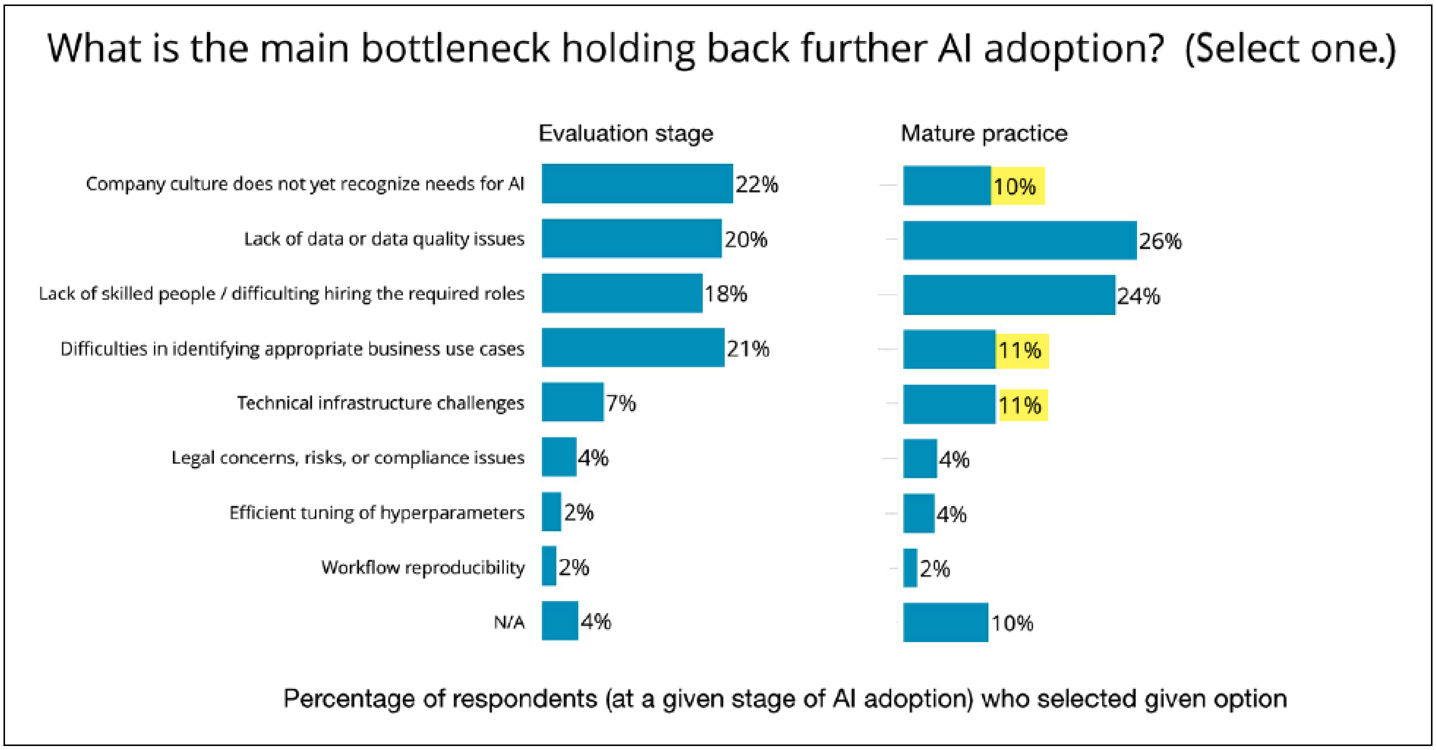


Figure 1–9. Challenges, by stage of maturity, in [4]

AI technologies for AI-Mature companies lean more to deep learning (highlighted in yellow), which require greater investments, as shown below.

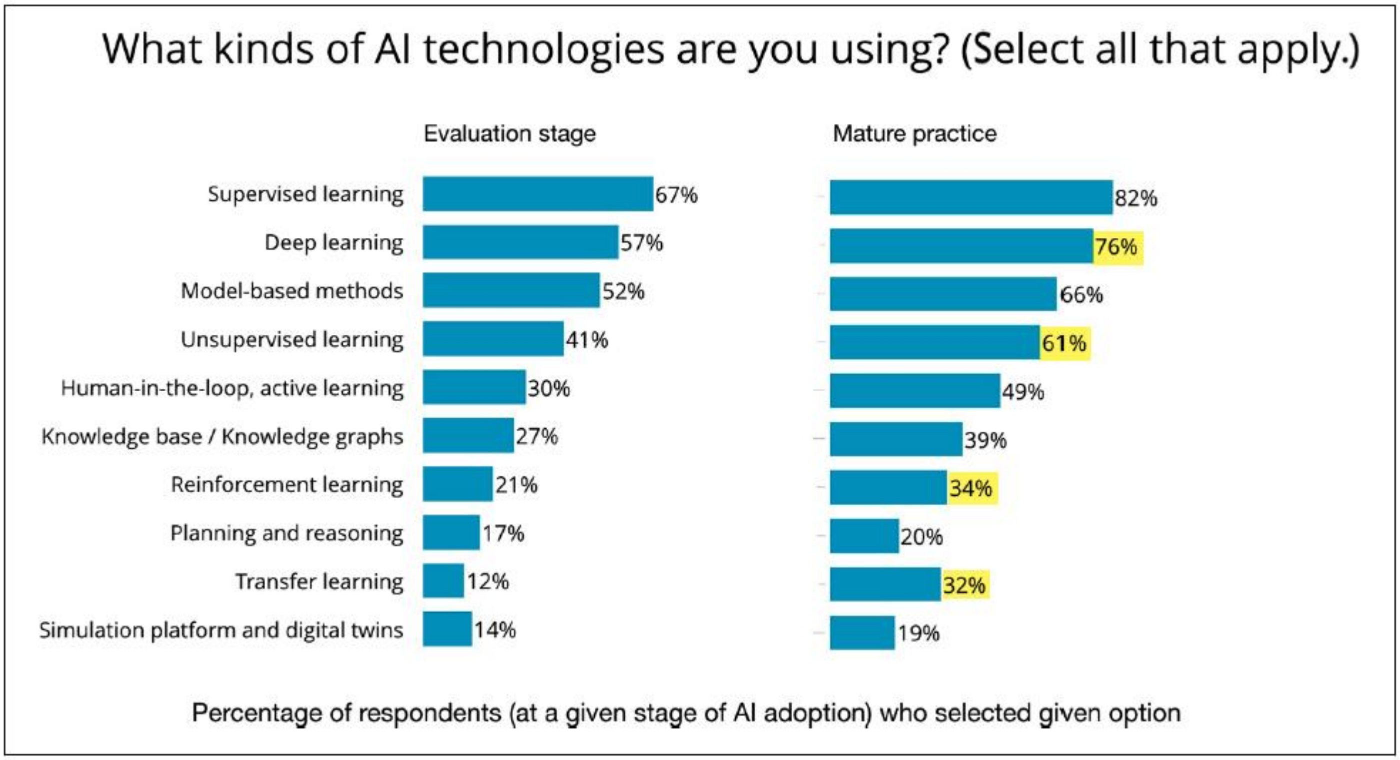


Figure 1–17. AI technologies, by stage of maturity, in [4]

The reason of the widening gap is that the barriers to AI adoption are mainly cultural, rather than technological. In this situation, there are no leap-frogging silver-bullets available for the lagers, at least not on the horizon!

Take-Away: A Zoo of AI Learning

Over the past year, one fascinating aspect about deep learning has been the ever-expanding *zoo* of approaches, methods, perspectives, and practices of deep learning. EVERY WEEK on arXiv, PapersWithCode, DataElixir, or TwoMinutePapers, I often react with a REALLY! …or AMAZING!

In the survey, you will see new terms for this *zoo*, like transfer learning, reinforcement learning, generative adversarial learning, and so on. One part of this *zoo* is a category called **meta-learning**, which denotes an algorithm that is learning about how it is learning and how to improve its learning. It is like the role of a human mentor who is guiding and provoking a student to learn more and better.

Refer above to Figure 1–17 on AI Technologies by stage of maturity. Note that Mature-Practice companies have 3x usage of transfer learning, as compares to Evaluation-Stage companies. Lorica & Nathan make this comment…

> *Transfer learning provides an interesting nuance, given how its use in production tends to require more experienced practitioners. We see mature practices making use of transfer learning at nearly three times the rate of evaluation stage companies. There's value in applications of transfer learning, although those [values] are perhaps not as apparent to the uninitiated. [4]*

Until companies research higher levels of AI maturity, they do not know what they do not know and are able to derive value.

Elsewhere Lorica & Nathan remark about reinforcement learning…

> *Respondents who already use reinforcement learning are beginning to build AI systems in some of the application areas for reinforcement learning that we listed in 2017: customer service; operations, facilities, and fleet management; finance; and marketing, advertising, and PR. [4]*

Finally, unsupervised deep learning is a HOT research area, with autoencoders, high-dimensional embedding spaces and the like. Many of the clear distinctions from machine learning are becoming muddy, such as the meaning of training/test datasets. Hence, absorbing this yet-another paradigm shift will be challenging!

Take-Away: Optimizing for Business Metrics — YES

Integral to culture change for successful AI systems is the skill of properly identifying business use cases for deep learning, shaped by metrics that target specific business objectives. This skill requires a rethinking of basic paradigms, plenty of imagination, and good business savvy. The lack of this skill across all functions (data scientists, data engineering, devOp specialists, LOB managers, etc.) is a critical roadblock.

In the figure below, the biggest skill gap (57%) was the obvious need for "ML modeler and data scientists". However, almost half (47%) indicates the requirement for people who can "understand and maintain a set of business use cases".

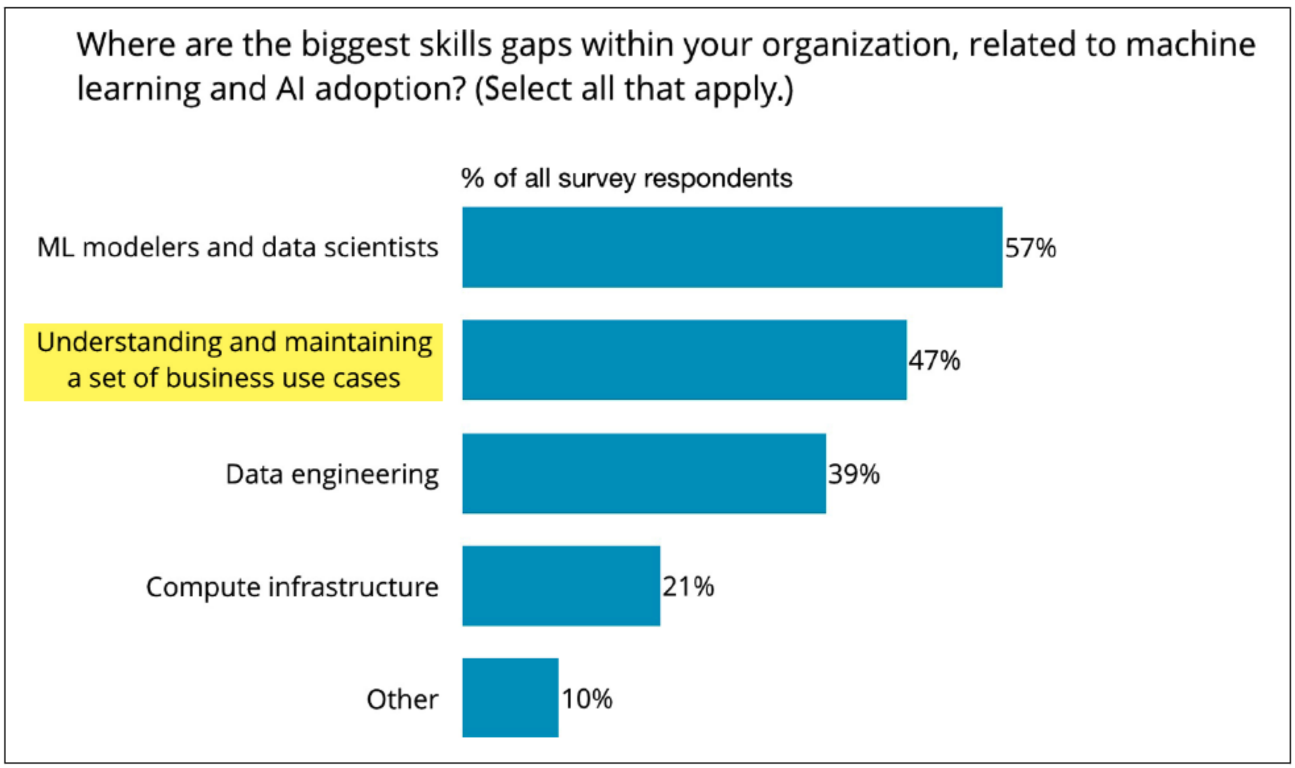


Figure 1–10. Skills gap, in [4]

Nathan specifically notes…

> *There's a dearth of proven recipes yet for "Product Management for AI," and that expertise will take time to cultivate at the organizational level. [1]*

Take-Away: Not Just About Business Metrics

One refreshing aspect is the emphasize on metrics that are not just about the business. Every organization exists within the large society, to which it must contribute, not just take. Hence, the recognition of survey respondents to the importance of “transparency, interpretability, fairness, bias, ethics, privacy, security, reliability, and compliance” was amazing.

The figure below lists the issues from the survey, separated by the AI Adoption maturity.

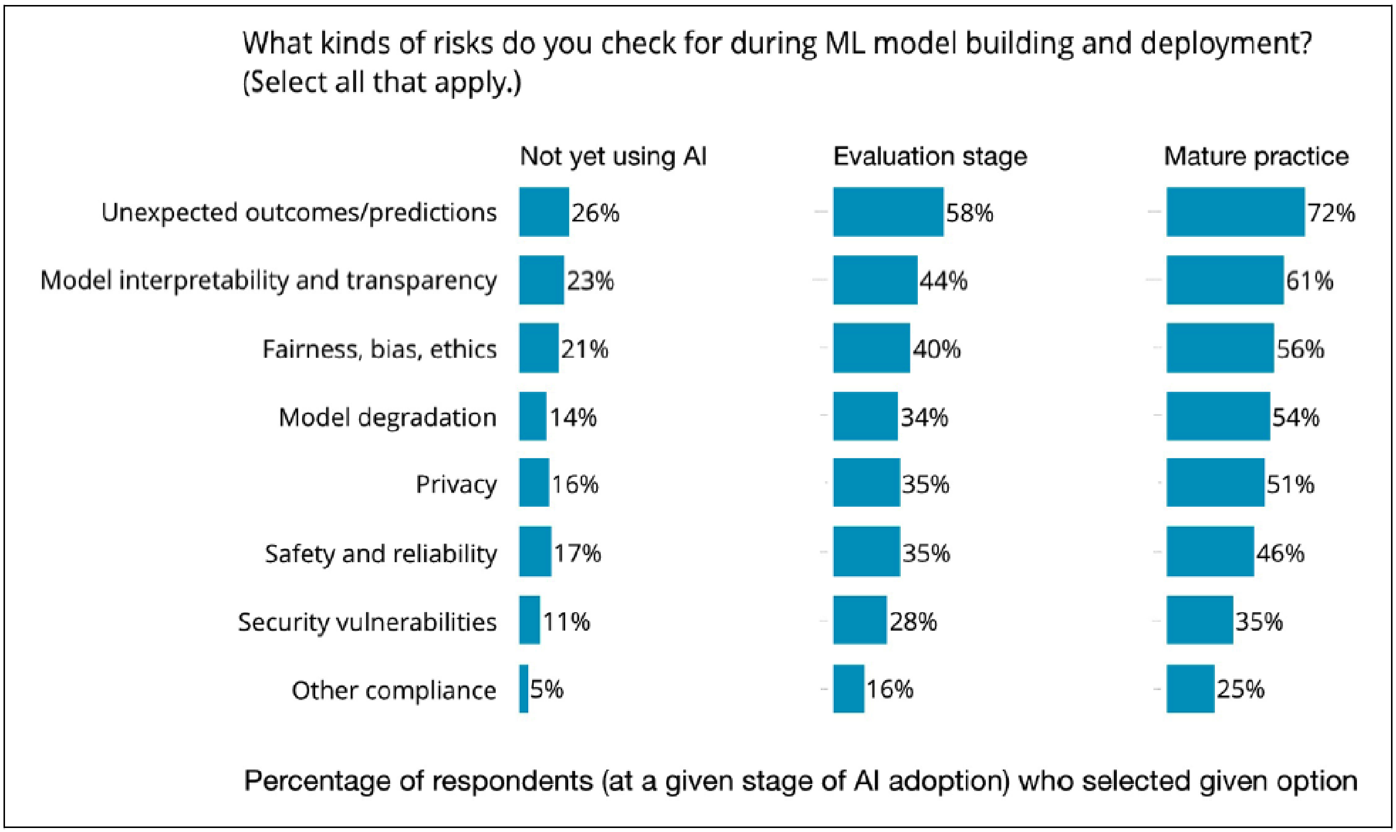


Figure 1–28. Risks checked, by stage of maturity, in [4]

Note that, as a company matures in its AI adoption, their concern for these issues increase. I detect a growing consensus among Data Science professionals that these issues are real risks to their company and will determine the long-term success of applying AI technology. There is hard technical work to be performed to make this a reality. Think of it like putting safety mechanisms on the nuclear bombs that you are manufacturing. This forces you into right frame of mind!

Take-Away: One Other Thing

Toward the end of Pacoid’s blog, there is a paragraph that starts “one other thing”, as if Nathan pauses, reflects, and shares a final deep concern…

> *A point that I keep hearing preached is how there’s a looming disconnect between data scientists and engineers. Instead, the looming disconnect which you need to stay up late worrying about is at the gap between engineering deployment of ML and the “last mile” of business use cases. Spark, Kafka, TensorFlow, Snowflake, etc., will not save you there. AutoML will not save you there. That’s the point*

where models degrade once exposed to live customer data, and where it requires significant statistical expertise to answer even a simple "Why?" question from stakeholders. That's the point where a large attack surface gets exposed to security exploits against the input data — with currently almost unimaginable magnitude of consequences. That's where complex ethics and compliance issues arise which cause angry regulators to come knocking. Those are business issues. Stop bloating data engineering teams as a panacea when ultimately the business issues are what will cause your organization the most grief. We're not enough years into AI adoption in enterprise yet for case studies about those issues to become standard HBS lectures, but they will be. Soon. [1]

This paragraph really hit me in the gut! I can clearly envision a large company enthusiastically launching an innovative AI system to solve a major business problem, only to crash-and-burn into the reality of the above paragraph.

Thinking New Paradigms and New Values

One useful way to comprehend the above take-away points is to realize that **WE ALL** need to think differently about the basic paradigms and values that have driven the technological advances of enterprise systems over past decades. This is not easy! Because…

It is a new way of thinking about **information** and how to utilize information effectively to improve business processes. Only generating insights for managers is no longer sufficient, which threats most of current Business Intelligence infrastructure today. Insights must not languish in committee meetings but must be immediately transformed into guidance for conducting today's business transactions.

It is a new way of thinking about **software**, which today is essentially art works of talented computer programmers …like paintings in a museum that portray cleaver pieces of static logic. The world is not static. Every piece of static logic is obsolete the day it goes into production, requiring constant hacking and patching to track the every-changing world. Software today fails to track these hourly zigzags of the world.

It is a new way of thinking about **learning**, which today is confined to data integration of an organization's information ecosystem, as reflected in data warehouse + data lake architectures. This paradigm must shift to managing models that generalized this data into a constant learning process. Software must track each zigzag of the world by constantly evolving the logic within those models. So, that AI system that you installed last week is now quite different in behavior this week. Can you manage that?

It is a new way of thinking about **human intelligence** and its role in managing the organization. It is becoming easier to think that **replacing** human intelligence with artificial intelligence is more cost efficient, hence it naturally leads to job elimination. One should accept this outcome and discover a new vocation. It is becoming harder to think that **augmenting** human intelligence with artificial intelligence (e.g., human-in-the-loop) has unique benefits that should be engineered into AI systems. This seems like an increasingly difficult task, without obvious benefits to the organization. So, is this the eventual outcome for human intelligence in future society?

Unfortunately, many organizations will never be able to transform their culture, thus losing their competitive advantage — the purpose for being a business — and suffering a lingering demise.

This is not inevitable! As IT professionals, we live in interesting times. Cough! With the exploding availability of AI technologies, we have exciting beneficial opportunities to transform businesses globally in ways that are now unimaginable. The hope is that, as a society, we do so to benefit all. However, the unfortunate reality is… If we are lazy or self-centered or even malicious, those same opportunities are available for misuse.

Plug for BizSmartAnalytics

I would like to collaborate with others who are concerned about the above issues. First, share your thoughts in the comments below. Second, share your feedback on related articles listed at BizSmartAnalytics.com. Third, support the Patreon to mentor peer group of IT professionals involved with AI systems at https://www.patreon.com/BizSmartAnalytics.

Tip of the hat to Lon Riesberg, editor of **Data Elixir** — which I recommend. Every week, I find at least one or two excellent resources about this exploding Data Science field. And, this is where I discovered the Pacoid blog. Subscribe for free at https://DataElixir.com/.

References

[1] Nathan, Paco. *Themes and Conferences per Pacoid, Episode 7*. Data Science Blog, Domino Data Lab. March 3, 2019.
https://blog.dominodatalab.com/themes-and-conferences-per-pacoid-episode-7/

[2] Lorica & Nathan. *The State of Machine Learning Adoption in the Enterprise.* August 2018.
https://www.oreilly.com/data/free/state-of-machine-learning-adoption-in-the-enterprise.csp (requires free registration w/o ad-blocking)

[3] Lorica & Nathan. *Evolving Data Infrastructure: Tools and Best Practices for Advanced Analytics and AI.* January 2019.
https://www.oreilly.com/data/free/evolving-data-infrastructure.csp (requires free registration w/o ad-blocking)

[4] Lorica & Nathan. *AI Adoption in the Enterprise: How Companies Are Planning and Prioritizing AI Projects in Practice.* February 2019.
https://www.oreilly.com/data/free/ai-adoption-in-the-enterprise.csp (requires free registration w/o ad-blocking)

[5] Lorica. *Three promising areas for AI skills development*. August 2018.
https://www.oreilly.com/ideas/3-promising-areas-for-ai-skills-development

[6] Lorica & Loukides. *How Companies Are Putting AI to Work Through Deep Learning*. April 2018.
https://www.oreilly.com/data/free/how-companies-are-putting-aI-to-work-through-deep-learning.csp

Link: https://towardsdatascience.com/confronting-deep-learning-systems-how-much-things-have-changed-f067738b728f